



```

67  rleuTyTrAspAspCysSerGlyProGlyGlnAspThrAspCysArgGluC 84
|||||
493  CTGTGACATGATCTCTGAGAGCGGAGAGATACAGAGATCTGAGGAGAT 442
84  ysluSerGlySerPheThrAlaSerGluAsnHisLeuArgHisCysLeu 100
|||||
443  GTGAGAGCGGCTGCTTACCGGCTCAGAAAAACCTCAGACACAGCGCTC 492
101  SerCysSerLysCysArgLysGluMetGlyGlnValGlnValGlnSerSerCy 117
|||||
493  AGCTGCTTCAAAATGCGAAAGCGGCGGAGAGAGAGAGAGAGAGAGAGAG 542
117  sThrValAspArgAspThrValCysGlyCysArgLysAsnGlnArgArg 134
|||||
543  CACAGTGTGACCGGAGACCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 592
134  tsTyTrpSerGluAsnLeuPheGlnCysPheAsnCysSerLeuCysLeu 150
|||||
593  ATATATCGAGTGAAGAAACCTTCTTCTGAGTGTCTCAATGCGAGCTGTG 642
151  AsnGlyThrValHisLeuSerCysGlnGlnLysGlnAsnThrValCysTh 167
|||||
643  AATGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 692
167  rCysHisAlaGlyPhePheLeuArgGlnAsnGlnCysValSerCysSerA 184
|||||
693  CTGCCATGCGAGTCTTCTTCTTAAAGAAAGAGAGAGAGAGAGAGAGAGAG 742
184  snCysLysLysSerLeuGlyCysThrLysLeuCysLeuProGlnLysGlu 200
|||||
743  ACHTAAAGAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 792
201  AsnValLysGlyThrGluAspSerGlyThrThrValLeuLeuProLeuVa 217
|||||
793  AATCTTAAAGAGATGACAGACATGAGGAGGAGGAGGAGGAGGAGGAGGAG 842
217  lLeuPheGlyCysCysSerGlnGlnSerGlnGlnGlnGlnGlnGlnGln 234
|||||
843  CATTTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 892
234  yTATTTTCGlnArgTrpLysSerSerSerSerSerSerSerSerSerSer 250
|||||
893  AIGCTTACAGACGGTGGAGAGCTTAAAGCTACCTGATTTGTGTGTGTGT 942
251  SerThrProGluLys 255
|||||
943  TCCACACCTTCAAAAA 957

```

seq\_name: gb\_est1:AU131978

```

seq_documentation_block:
Locus      AU131978      761 bp      mRNA      EST      24 cont 2000
DEFINITION AU131978 NT2RP3 Homo sapiens cDNA clone NT2RP3003570 5', mRNA
sequence
ACCESSION  AU131978
VERSION    AU131978.1 GI:10992332
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini, Hominoidea; Homo.
REFERENCE  1 (bases 1 to 761)
AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
            Yamamoto,J., Wakamatsu,A., Nakamura,Y., Naga,T., Sugano,S. and
            Isogai,T.
            HRI human cDNA project
            Unpublished (2000)
            Contact: Takao Isogai
            Genomics Laboratory
            Helix Research Institute
            1542-4 Yana, Kisarazu, Chiba 292 0812, Japan
            tel: 81-438-52-3951

```

Fax: 81-438-52-3952

Email: genomics@hri.co.jp

HRI human cDNA project; 5' & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

# FEATURES

## SOURCE

1..761

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="NT2RP3003570"

/clone\_lib="NT2RP3"

/cell\_type="teratocarcinoma"

/cell\_line="NT2"

/note="Vector: pME18SL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"

BASE COUNT 178 a 199 c 201 g 179 t 4 others

## ORIGIN

## alignment\_scores:

Quality: 1268.00 Length: 234

Ratio: 5.489 Gaps: 1

Percent Similarity: 98.718 Percent Identity: 97.863

## alignment\_block:

US-09-525-998A-2 x AU131978

Align seq 1/1 to: AU131978 from: 1 to: 761

```

1  MetGlyLeuSerThrValProAspLeuLeuLeuProLeuValLeuLeuG 17
|||||
62  ATGGAGATCTCTGAGAGCGGAGAGATACAGAGATCTGAGGAGATCTGAG 111
17  rLeuLeuValGlyHisTyProSerGlyValHisGlyLeuValProHisL 34
|||||
112  CCGTGTGCGGGGAATATACCGCTTACGGGGTTATGGACTGGTGGCTCACC 161
34  euGlyAspArgGluLysArgAspSerValCysProGlnGlyLysTyrlle 50
|||||
162  TAGGCGTAGGCGAGAGAGATATGCTGTCTGTCTGTCTGTCTGTCTGTCT 211
51  HisProGlnAsnAsnSerLysCysCysThrLysCysHisLysGlyThrly 67
|||||
212  CACGCTCAAAAATAATTCGATTTGTTGTTAGTAAAGTGGCACAAGAGAGCTA 261
67  rLeuTyTrpAsnAspCysProGlyGlnAspThrAspCysArgGluC 84
|||||
262  CTTGTACAAATACCTGTGTAGGCGCGGAGAGAGATACCGAGCTGCAAGGAG 311
84  ysluSerGlySerPheThrAlaSerGluAsnHisLeuArgHisCysLeu 100
|||||
312  GTGAGAGCGGCTGCTTACCGGCTCAGAAAAACCTCAGACACAGCGCTC 361
101  SerCysSerLysCysArgLysGluMetGlyGlnValGlnValGlnSerSer 117
|||||
362  AGTGTGTCTTAAATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 411
117  sThrValAspArgAspThrValCysGlyCysArgLysAsnGlnArgArg 134
|||||
412  CACATGAGACCTGGGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 461
134  tsTyTrpSerGluAsnLeuPheGlnCysPheAsnCysSerLeuCysLeu 150
|||||
462  ATTATTGTAGTCAAAAATTTTTCAGTGTCTTAAATGCAATGCAATGCTGCT 511
151  AsnGlyThrValHisLeuSerCysGlnGlnLysGlnAsnThrValCysTh 167
|||||
512  AATGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 561
167  rCysHisAlaGlyPhePheLeuArgGlnAsnGlnCysValSerCysSerA 184
|||||
562  CTGCCATGCGAGTCTTCTTCTTAAAGAAAGAGAGAGAGAGAGAGAGAGAG 611

```

184 seqsLysLysSerLeuGlyThrLysLeuGlyLeuProGlnIleGlu 200  
 |||||  
 612 ACTGTAAACAAAGCGTGGAGTGCACGAAGTGGTGGTACCCACAGATTGAC 661  
 |||||  
 201 AsnValLysGlyThrGluAspSerGlyThrValLeuLeuLeuProLeuV 217  
 |||||  
 662 AATGTTAANAGCATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 711  
 |||||  
 217 atttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 233  
 |||||  
 712 NCATTTCTTGGTCTTGGTCTTGGTCTTGGTCTTGGTCTTGGTCTTGGTCTT 761  
 |||||

seq\_name: qb\_est1:AL529836

seq\_documentation\_block: 942 bp mRNA EST 13 FEB-2001  
 LOCUS AL529836 L1L\_NFL001\_NBC4 Homo sapiens cDNA clone CSDD005YP05 5  
 DEFINITION prime, mRNA sequence  
 ACCESSION AL529836  
 VERSION AL529836.1 GI:12793129  
 KEYWORDS EST  
 SOURCE human  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata, Vertebrata, Euteleostomi,  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 942)  
 AUTHORS Li, W.-H., Gruber, C., Jessup, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

FEATURES  
 1..942  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /catalytic="catalytic"  
 /catalytic="catalytic"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="H10N10"  
 /note="organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9600 Medical Center Drive, Foster City, CA 94024, USA Fax: (1) 303 519 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 208 a 259 c 248 g 224 t 3 others  
 ORIGIN

alignment\_scores:  
 quality: 1261.00 length: 237  
 Ratio: 5.389 Gaps: 1  
 Percent Similarity: 98.734 Percent Identity: 97.890

alignment\_block:  
 US-09-525-998a-2 x AL529836 ..  
 Align seq 1/1 to: AL529836 from: 1 to: 942

1 MetGlyLeuSerThrValProAspLeuLeuLeuProLeuValLeuLeu 17  
 |||||  
 234 ATGGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 283  
 |||||  
 17 uLeuLeuValGlyLeuValProSerGlyValIleGlyLeuValProIle 34  
 |||||

284 GCTGTTGTTGGGAAATATATACGCTTAAAGGATATATATATATATATATATAT 343  
 |||||  
 34 ouGlyAspPheGlyAspSerValGlyCysProGlnIleGlyLysPhe 50  
 |||||  
 334 TAGGGGACATGACACAGAGATAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 483  
 |||||  
 51 HisProGluAsnAspSerLeuGlySerValThrLysCysHisLysValThr 67  
 |||||  
 484 CACCTTCACAAAATATATATATATATATATATATATATATATATATATATATAT 443  
 |||||  
 67 tLeuValAspAspGlyProGlyThrValLeuAspThrAspGlyLeu 84  
 |||||  
 434 CTGTACAAATATATATATATATATATATATATATATATATATATATATATATAT 483  
 |||||  
 84 YsGluSerGlySerProThrAlaSerGlnAspHisLeuValThrHisCysLeu 1000  
 |||||  
 484 GTGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 533  
 |||||  
 101 SerCysSerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 117  
 |||||  
 534 AAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 583  
 |||||  
 117 sLeuValAspAspValGlyValGlyCysGlyValGlyValGlyValGlyValGly 134  
 |||||  
 584 CAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 633  
 |||||  
 134 tStyTrpSerGlnAsnLeuPheGlnCysGlnAspCysSerLeuCysLeu 150  
 |||||  
 634 ATATATGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 683  
 |||||  
 151 AsnGlyThrValHisLeuSerCysGlnHisLeuValGlyValGlyValGlyVal 167  
 |||||  
 684 AATGGGACCGCTGAAATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 733  
 |||||  
 167 tCysHisAlaGlyPhePheLeuAspGlnAsnLeuCysValSerCysSer 183  
 |||||  
 734 CTGGCATGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 783  
 |||||  
 184 AsnCysLysLysSerLeuGlnCysThrLysCysGlySerGlnHisLeuValGly 200  
 |||||  
 784 AACGTAAAGTAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 832  
 |||||  
 200 uAsnValLysGlyThrValAspSerGlyThrValValGlyLeuLeuLeuVal 217  
 |||||  
 832 CAAATGTTAAAGTAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 882  
 |||||  
 217 atttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 233  
 |||||  
 882 tCAlIITCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 932  
 |||||  
 233 tTGAAGTGT 246  
 |||||  
 933 TATGCTTAC 941

seq\_name: qb\_est1:AL577008

seq\_documentation\_block:

LOCUS AL577008 975 bp mRNA EST 16-FEB-2001  
 DEFINITION prime, cDNA sequence.

ACCESSION AL577008  
 VERSION AL577008.1 GI:12939716  
 KEYWORDS EST  
 SOURCE human

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 975)  
 AUTHORS Li, W.-H., Gruber, C., Jessup, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage

BP 191 91006.EMVY.cds.1 - France	
Email : seqrel.genoscope.cns.fr, Web : www.genoscope.cns.fr.	
location/Qualifiers	
1. 975	
"organism"="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="GS051082XA01"	
/clone_lib="LT1_NFL005_PL2"	
/assay_type="parenta"	
/vector="pMYSKnot 6, Site1, NotI, 1st strand cDNA	
was primed with a NotI-oligo(dN) primer. Five prime end	
enriched, double-stranded cDNA was digested with NotI and	
cloned into the NotI and EcoRV sites of the pMYSKnot 6	
vector. Library was normalized. Library was constructed by	
life technologies. Contact : Food and life technologies,	
a division of invitrogen 9800 Medical Center Drive	
Rockville Maryland 20850, USA Fax : (1) 301 430 8371	
Email : lilian@litech.com URL :	
http://lilientech.invitrogen.com	
231 a 258 c 259 g 221 t 6 others	
BASE COUNT	
ORIGIN	

[illegible]

167	TCYSHLSALUGLYPFPheLeuArgGluAlaGluCysValSerCysSerA	184
756	CTCCCATGCCAGTTTCTTCTTCAAGAGAAAACGACCTGTCTCTCTCTACTA	805
184	SNCGSYLSYSSYSSerLeuGluCysThrTyrSLeu.CysLeuLeuGlnLLeu	200
806	ATCTGATGAAAAGTGTGATGTAAGATGTTTGTATGCTTATGCTGATG	855
200	AGSNValTyrGlyThrGluAspSerGlyThrValTyrCuleuThrSLeuV	217
856	GAAGCTTAAGGCGATGAGGATCTACGATGATGAGTGGGCTGTGGCCCTGG	905
217	allLeuPheGlyLeuGlySLeu.LeuSerLeuLeuPheLeuGlyLeuMe	233
906	TCACTTCTTGGGCTGCTTGGCTTATGCGG...CTCTGCAATGGCTTAA	952
233	TTTATGTTTTCGln	237
953	GTATGCTGATGAA	965
seq_name	qt_est2:HG180101	
seq_documentation_block:		
LOCUS	BG180101 974 bp mRNA	EST
DEFINITION	602324674F1 NHL_M3_91 Homo sapiens cDNA clone IMAGE 4441819 5'	06-FEB-2001
ACCESSION	BG180101	
VERSION	BG180101.1	
KEYWORDS	EST	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1 (bases 1 to 974)	
TITLE	NHL-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cdapts-remail.nih.gov">cdapts-remail.nih.gov</a> Tissue Procurement: DCTD/PPP cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The 2-M.A.G.E. Consortium (LLNL) cDNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the 1 M.A.G.E. Consortium/HLNI at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> plate: LLAM10185 row: g column: 20 High quality sequence stop: 657.	

```

FEATURES              source
high quality sequence stop : 607
Location/Qualifiers
1 . 974
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone-"IMAGE:4431019"
/clone_lib="NIL_MGC_91"
/tissue_type="adenocarcinoma, cell_line"
/lab_host="DR10B (phage-resistant)"
/note="Organ: prostate; Vector: pCMV-SpeK16; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: This is a NIL MGC Library."
2754      2853     2854     2855     2856     2857     2858     2859
BASE COUNT
        G       C       A       T
ORIGIN

```

```

alignment_scores:
  Quality: 1234.00
  Ratio: 5.412
  Percent similarity: 98.272
  Length: 232
  Gaps: 1
  Percent identity: 97.845

alignment_block:
  05-09-525-998A-2 x BG180101

```

Align seq 1/1 to: HG180101 from: 1 to: 974

```

2 GlyLeuSerThrValProAspLeuLeuLeuProLeuValLeuLeuGluLeu 18
|||||
1 GGGTCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 50
|||||
18 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 35
|||||
51 GTTGGTGGGAAATATAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 100
|||||
35 LysAspArgGluLeuAspSerValGlyProGlyProGlyProGlyProGly 51
|||||
101 GGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 150
|||||
52 ProGluAsnAsnSerThrValGlyProGlyProGlyProGlyProGlyPro 58
|||||
151 GCTCAAAATATTCGATTTCGTTGATTAAGTACACCAAGGAGGAGGAGGAGT 200
|||||
68 nryASnAspCysProGlyProGlyProGlyProGlyProGlyProGlyPro 95
|||||
201 GTCATATGATGTCATGATGATGATGATGATGATGATGATGATGATGATGAT 250
|||||
85 LysGlySerProThrAlaSerGluAsnGluLeuAlaGlyCysLeuSer 101
|||||
251 AGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
|||||
102 CysSerGlyCysArgLysGluMetGlyGluValGluLeuSerSerCysThr 118
|||||
301 TGTCTCAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 350
|||||
118 rValAspArgAspThrValGlyCysArgLysAsnGlnIlyrArgHis 135
|||||
351 AGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 400
|||||
135 YrTrpSerGluAsnLeuPheThrLysPheAsnGlySerLeuLysLeuAsn 151
|||||
401 ATTCGACAGCAAAACCTTTTCACAGCTGCTTCAATTCGACGCTGCTTCAAT 450
|||||
152 GlyThrValHisLeuSerGlyGluLeuValGluLeuValGluLeuValGlu 169
|||||
451 GGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 500
|||||
168 SHSAlaGlyPheLeuLeuArgGluAsnGluCysValSerCysSerAspG 185
|||||
501 CGATGAGATTTTCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 550
|||||
185 YsLysLysSerLeuGluCysThrLysLeuCysLeuProGlnIleGluAsn 201
|||||
551 GTAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
|||||
202 ValLysGlyThrValAspSerGlyThrThrValLeuLeuProLeuValGlu 218
|||||
601 GTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 649
|||||
218 ePheGlyGlyLeuCysLeuSerLeuSerLeuPheLeuGlyLeuGlyLeu 232
|||||
650 TTCTTTGGTCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCT 693
|||||

```

seq\_name: qb.esl1:A0117362

```

seq documentation block:
LOCUS      A0117362      889 bp      mRNA      ESI      19-OCT-2000
DEFINITION A0117362 HEMERAL Homo sapiens cDNA clone HEMBAL001229 5', mRNA
sequence.
ACCESSION  A0117362
VERSION    A0117362.1
KEYWORDS   EST, GI:10942324
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 889)
AUTHORS   Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,

```

Yamamoto,T., Wakamatsu,A., Nakamura,Y., Naito,T., Souda,H.S. and Isozaki,T.

TITLE  
JOURNAL  
COMMENT

HRI human cDNA project  
Unpublished (2000)  
Contact: Takao Isozaki  
Genomics Laboratory  
Helix Research Institute  
1532-4 Yada, Kiserazu, Chiba 292 0812, Japan  
Tel: 81 438-52-4551  
Fax: 81 438-52-4952  
Email: genomics@hri.co.jp

HRI human cDNA project; 5'- 6' end one pass sequencing; Helix  
Research Institute; cDNA library construction; Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

FEATURES

Source

1..889  
Location/Qualifiers  
/organism="Homo sapiens"  
/library="cDNA project"  
/clone="HEMBAL001229"  
/clone\_lib="HEMBAL"  
/issue\_type="Whole embryo, mainly head"  
/map\_to="embryo, 10 weeks"  
/note="Vector: pMIRSP13"

BASE COUNT 107 a 301 c 249 g 166 t 6 others  
ORIGIN

alignment\_scores:

Quality: 1206.00 Location: 229  
Ratio: 5.414 Gaps: 0  
Percent similarity: 59.127 Percent identity: 98.640

alignment\_block:

US-09-525-998A.2 x A0117362

Align seq 1/1 to: A0117362 from: 1 to: 889

```

223 CysLeuLeuGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 439
|||||
4 TGGCTTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 52
|||||
239 PysSerLysSerLeuGluLeuValGlySerGlySerGlySerGlySerGlySer 256
|||||
54 CAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 102
|||||
256 LysLysGluLeuGluGluLeuThrThrThrThrThrThrThrThrThrThr 272
|||||
104 ACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 352
|||||
273 PheSerProThrProGlyPheThrThrThrThrThrThrThrThrThrThr 289
|||||
153 TTGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 292
|||||
282 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 309
|||||
204 CAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 352
|||||
306 subPheAlaAlaProAlaAlaValAlaValAlaProProProProProPro 322
|||||
253 ACCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 382
|||||
323 ProLysLeuAlaThrAlaAlaAlaSerAspProLysProLysProLysPro 349
|||||
303 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 352
|||||
339 nryLysLysLysSerAlaHisLysProGluSerLeuAspSerLeuAspSer 356
|||||
353 GAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 402
|||||
356 rValAlaThrLeuThrValAlaValGluAsnValProProLeuAspThrLys 372
|||||
404 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 452
|||||

```









```

309 aProArqAqGluValAlaProProTyrGlnGlyAlaAspProIleLeuA 326
|||||
201 TCCCGGACAGAGGTGGCAGCAGGCTATCAGCGGGCTCAGCCCATCTTTG 250
|||||
326 GATrAlaLeuAlaSerAspProIleProAsnProLeuGlnGlyIlePhe 342
|||||
251 GCAAGAGCTCGGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 300
|||||
343 AspSerAlaHisGlySerGlnSerLeuAspThrAspProAlaThrIle 359
|||||
301 GACAGCGGCCACACGCCAGCAGCCTACACACTCATCAGCCCGGAGCCT 350
|||||
359 uTyAlaValValGluAsnValProProLeuAqITrPlyGluPheValA 376
|||||
351 GACCGCGCTGGTGGAGAACCGTGGCGCGCTGGCGTGGAGAACCAATTC 400
|||||
376 rTArqGluGlyGluSerAspSerGlyLeuAspArqGluGlnGlnAsp 392
|||||
401 GCGGCTAGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 450
|||||
393 GlyArqGlySerLeuArqGluAlaGlnIleTyrSerMetLeuAlaThrTrpArqAr 409
|||||
451 GCGGCTGGCTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 500
|||||
409 qArqThrProArqAqGluAlaThrLeuGlnLeuLeuGlyArqValLeuA 426
|||||
501 GCGACGCTGGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGG 550
|||||
426 rTArqPheTAspLeuGlnGlySerGlnGlyAspIleGlnGlnAlaLeuGly 442
|||||
551 GCGACATGAGGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATG 600
|||||
443 GlyProAlaAlaLeuProProAlaPro 451
|||||
601 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 527

```

seq\_name: qb\_est2-100282718

```

seq_documentation_block:
  LOCUS      602406578P1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE 4518690 c.
  DEFINITION mRNA sequence.
  ACCESSION  BC282718
  VERSION    BC282718.1 GI:113031865
  KEYWORDS   EST.
  SOURCE     human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE  NIH-MGC http://mgi.mit.edu/
             National Institutes of Health, Mammalian Gene Collection (MGC)
             Unpublished (1999)
  AUTHORS    Contact: Robert Strausberg, Ph.D.
             Email: cga@pilot.mcg.edu
             Tissue Procurement: DCTD/DP
             cDNA library preparation: Life Technologies, Inc.
             cDNA library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: L1AM10413 row 1 column: 19
             High quality sequence stop: 682.
  FEATURES   location/coordinates
             1..761
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="IMAGE:4518690"
             /clone_lib="NIH_MGC_91"
             /tissue_type="adipocarcinoma, cell line"
             /lab_host="PH10B (phage-resistant)"
             /note="organ: prostate, Vector: p.MV SP6RT6; Site:1; Note:

```

```

Site_2: Salt: Cloned unidirectionally; oligo dT primed.
Average insert size 1.4 kb. Library enriched for
500 length clones and constructed by life technologies.
Note: this is a NIH_MGC library.
BASE COUNT  176 G  228 C  179 A  179 T
ORIGIN

```

```

alignment_statistics:
  Quality: 110.50      Length: 285
  Ratio: 4.646        Gaps: 11
  Percent Similarity: 83.860  Percent Identity: 78.596
alignment_blocks:
  US-09-525-998a-2 x BC282718
  Align: 503 173 to 60282718 from 1 to 761
  61  LysGlyHisGlyGlyThrTyrLeuTyrAsnAspGlyGlyGlyGlyGly 77
  |||||
  2  AAGTGGCAGCAAGAGAACTACTTGTACAAAGAGTGTGTGTGTGTGTGT 51
  |||||
  77  GAspThrAspGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 93
  |||||
  52  GGAATAGGATTAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 101
  |||||
  94  AsnHisGlnGlnHisGlySerGlyGlyGlyGlyGlyGlyGlyGlyGly 110
  |||||
  102  AAGCAGCTGAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 151
  |||||
  110  LysLeuAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 126
  |||||
  152  GTCAGGTGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 201
  |||||
  127  CysArqGlyAsnGlnIleTyrArqHisTyrTyrSerGluAsnLeuPheGly 143
  |||||
  202  TGTAAAGAACAAAGATAGGAGATATATAGAGAGAGAGAGAGAGAGAG 251
  |||||
  143  sPrAsnGlySerHisGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 159
  |||||
  252  GTTAAATGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 301
  |||||
  160  LysLeuGlnAsnThrValGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 176
  |||||
  302  AGAAGCAGAAATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 451
  |||||
  177  AsnHisGlyValSerGlySerGlySerGlySerGlySerGlySerGlySer 193
  |||||
  252  AAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 276
  |||||
  193  LysGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 209
  |||||
  376  .....
  210  ThrThrValLeuLeuProLeuValIleLeuPheGlyGlyGlyGlyGlyGly 226
  |||||
  377  ACCACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 426
  |||||
  226  GlnHisGlnHisGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 242
  |||||
  427  GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 476
  |||||
  243  LeuTyrSerThrValGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 259
  |||||
  477  GTTACTGCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 526
  |||||
  259  GlnGlyThrThrThrLeuLeuProLeuValIleLeuPheGlyGlyGlyGly 276
  |||||
  527  TGAAGGAGACACACACACACACACACACACACACACACACACACACAC 576
  |||||
  276  ThrThrGlyThrThrThrThrThrThrThrThrThrThrThrThrThrThr 292
  |||||
  577  AATGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626

```





```

51 115PProGlnAsnAsnSerIleCysCysThrLysCysHisLysGlyThrTy 67
|||||
370 CACCCCTCAAAATAATTCGATTTCGCTGTAACAAGTCCACAAAGCAACCTA 419
|||||
67 115PProGlnAsnAsnSerIleCysCysThrLysCysHisLysGlyThrTy 84
|||||
420 GTTGTATCAATGACTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 459
|||||
84 115PProGlnAsnAsnSerIleCysCysThrLysCysHisLysGlyThrTy 100
|||||
470 GTTGTATCAATGACTGTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 519
|||||
101 SerCysSerLysCysArgLysGlnMetGlyGlnValGluLleSerSerCy 117
|||||
520 AGCTGGTCAAAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 569
|||||
117 115PProGlnAsnAsnSerIleCysCysThrLysCysHisLysGlyThrTy 134
|||||
570 CACAGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 619
|||||
134 115PProGlnAsnAsnSerIleCysCysThrLysCysHisLysGlyThrTy 150
|||||
620 ATTATTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669
|||||
151 115PProGlnAsnAsnSerIleCysCysThrLysCysHisLysGlyThrTy 167
|||||
670 AATCGGACCGTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
|||||
167 115PProGlnAsnAsnSerIleCysCysThrLysCysHisLysGlyThrTy 184
|||||
720 CTCTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
|||||
184 115PProGlnAsnAsnSerIleCysCysThrLysCysHisLysGlyThrTy 200
|||||
769 ACTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 817
|||||
201 AsnValLysGlyThrGluAsp 207
|||||
818 ...TGTTAAGGACATGAGGAC 835

seq_name: qb_est1:A0124446
seq_documentation_block:
LOCUS       A0124446             859 bp     mRNA             23-OCT-2000
DEFINITION A0124446 NT2PM4 Homo sapiens cDNA clone NT2PM400018 5', mRNA
sequence.
ACCESSION   A0124446
VERSION     A0124446.1
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 859)
AUTHORS     Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S., Salto,K., Yamamoto,J.,
            Nakamura,Y., Nishikawa,T., Nagai,T., Soraki,Y., Sugano,S. and
            Isoqai,T.
TITLE       HRI human cDNA project (Ota,T., Wakamatsu,A., Ozawa,M., Ishii,S.,
            Salto,K., Yamamoto,J., Nakamura,Y., Nishikawa,T., Nagai,T., Suzuki
            Y., Sugano,S., Isoqai,T.)
JOURNAL     Unpublished (2000)
COMMENT     Contact: Takao Isoqai
            Genomics Laboratory
            Helix Research Institute
            1532-3 Yaba, Kisarazu, Chiba 292-0812, Japan
            Tel: 81-438-52-3951
            Fax: 81-438-52-3952
            Email: genomics@hri.co.jp
            HRI human cDNA project 5' & 3'-end one pass sequencing: Helix
            Research Institute cDNA Library Construction Department of
            Virology, Institute of Medical Science, University of Tokyo, and
            Helix Research Institute.

```

```

FEATURES             Location/Qualifiers
     source            1..859
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="NT2PM400018"
                     /clone_lib="NT2PM4"
                     /cell_type="teratocarcinoma"
                     /cell_line="NT2"
                     /note="vector: pME18SFL3; mRNA from uninduced N12 neuronal
                     precursor cells"
BASE COUNT          197 a   247 c   220 g   192 t   3 others
ORIGIN
alignment_scores:
    Quality: 1048.00      Length: 194
    Ratio: 5.458          Gaps: 0
    Percent Similarity: 98.969 Percent Identity: 98.454
alignment_block:
US-09-525-998A-2 x A0124446
Align seq 1/1 to: A0124446 from: 1 to: 859
1  MelGlyLeuSerThrValProAspLeuLeuProLeuValLeuLeuL 17
|||||
279  ATGGGCTCTCCACCGCTGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCT 328
|||||
17  leuLeuValGlyIleCtyProSerGlyValIleCtyLeuValProHisL 34
|||||
329  GTTGTGTGAGAGAAATATACTTAAAGAGATATATGAGATGAGTGGTGG 378
|||||
34  euCtyAspArgCulLysArgAspSerValCysProGlnCtyLysThrIle 50
|||||
379  TAGGGGACAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 428
|||||
51  HisProGlnAsnAsnSerIleCysCysThrLysCysHisLysGlyThrTy 67
|||||
429  CACCCCTCAAAATAATTCGATTTCGCTGCTGCTGCTGCTGCTGCTGCTG 478
|||||
67  ileuTyAsnAspCysProGlyProGlyClnAspThrAspCysArgGluC 84
|||||
479  CTGTGACAAATGATTTCTAAAGTGTGAGAGAGAGAGAGAGAGAGAGAG 528
|||||
84  ysClnSerGlySerIleThrAlaSerGluAsnHisLysGlyThrTy 100
|||||
529  GTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 578
|||||
101  SerCysSerLysCysArgLysGluMetGlyGlnValGluLleSerSerCy 117
|||||
579  AGCTGCTCAAAAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 628
|||||
117  sThrValAspArgAspThrValCysGlyCysArgLysGlnValGluArg 134
|||||
629  CACCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 678
|||||
134  isTyTrpSerGluAsnLeuPheGlnCysPheAsnGlySerLysCysLeu 150
|||||
679  ATTATTTGCACTCAAAAGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCT 728
|||||
151  AsnGlyThrValHisLeuSerCysGlnGlnGlyLysGlnAsnThrValCys 167
|||||
729  AATGGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
|||||
167  rCysHisAlaGlyPhePheLeuArgClnAsnGlnCtyLysValSerCys 184
|||||
778  CTCTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 827
|||||
184  snCysLysLysSerLeuGluCysThrLysLeu 194
|||||
828  ACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859

```



